



RAW SEQUENCE LISTING **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/868,328
Source: PCT09/
Date Processed by STIC: 6-20-02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebs/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Does Not Comply
Corrected Diskette Needed



PCT

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/868,328

DATE: 06/20/2002
TIME: 14:07:32

Input Set : A:\EP.txt
Output Set: N:\CRF3\06202002\I868328.raw

3 <110> APPLICANT: RHEE, Sangki
4 SONG, Kibang
5 KIM, Chulho
6 RYU, Eunja
7 Lee, Yongbok
9 <120> TITLE OF INVENTION: Enzymatic Production of Diffructose Dianhydride IV from
Sucrose and Relevant
10 Enzymes and Genes Coding for them
12 <130> FILE REFERENCE: 24679
14 <140> CURRENT APPLICATION NUMBER: US 09/868,328
16 <141> CURRENT FILING DATE: 2001-06-18
18 <150> PRIOR APPLICATION NUMBER: PCT/KR00/01183
20 <151> PRIOR FILING DATE: 2000-10-19
22 <160> NUMBER OF SEQ ID NOS: 4
24 <170> SOFTWARE: PatentIn 3.1

ERRORED SEQUENCES

66 <210> SEQ ID NO: 2
67 <211> LENGTH: 521
68 <212> TYPE: PRT
69 <213> ORGANISM: Arthrobacter ureafaciens K2032
W--> 70 <400> SEQUENCE: 2
71 MTPAISRRRAV LQAGAGALA LIFGGAVPPA ARASAPGSLR AVYHMTPPSG WLCDPQRPVT 60
72 THGAYQLYYL HSDQNGPGG WDHASTDGV AFTHHGTVM LRPDFPVWSG SAVVGTANTA 120
73 GFGAGAVVAL ATQPTDGVRK YQEYLYWST DGGFTFTALP DPVIVNTDGR AATTPAEIEN 180
74 AEWFRDPKIH WDTARGEWVC VIGRLRYAAF YTSNLRDWT LRRNFDYPNH ALGGIECPDL 240
75 FEITADDGTR HWVLAASMDA YGIGLPMTYA YWTGTWDGEQ FHADDLTPQW LDWGWDWYAA 300
76 VTWPSIDAPE HKRLAIAMWN NWKYAARDVP TDASDGYNGQ NSIVRELRLA RQPGGWYTL 360
77 STPVAALTNV VTATTTLPDR TVDGSVLPW NGRAYEIELD IAWDTATNVG ISVGRSPDGT 420
78 RHTNIGKYGA DLYVDRGPSD LAGYSLAPYS RAAAPIDPGA RSVHLRILVD TQSVEVFVNA 480
E--> 79 GHTVLSQQVH FAEGDTGISL YTDGGAHFT GIVVREIGQA I 521
117 <210> SEQ ID NO: 4
118 <211> LENGTH: 423
119 <212> TYPE: PRT
120 <213> ORGANISM: Zymomonas mobilis
W--> 121 <400> SEQUENCE: 4
122 MLNKAGIAEP SLWTRADAMK VHTDDPTATM PTIDYDFPVM TDKYWWDTW PLRDINGQVV 60
123 SFQGSVIFA LVADRTKYGW HNRNDGARIG YFYSRGSNW IFGGHLLKDG ANPRSEWWSG 120
124 CTIMAPGTAN SVEVFFTSVN DTPSESVPAQ CKGYIYADDK SVWFDGFDKV TDLFQADGLY 180
125 YADYAENNFV DFRDPHFVIT PKIGKTYALF EGNVAMERGT VAVGEEEEIGP VPPKTETPDG 240
126 ARYCAAAGI AQALNEARTE WKLLPPLVTA FGVDQTERP HVVFQNGLTY LFTISHHSTY 300
127 ADGLSGPDGV YGFVSENGIF GPYEPLNGSG LVLGNPSSQP YQAYSHYVMT NGLVTSFIDT 360

see CFR 1.822 (d)

must use 3 letter
codes, number every
5, and have a max-
imum 16 amino acids/line

see examples
attached to
end of report.

RAW SEQUENCE LISTING

DATE: 06/20/2002

PATENT APPLICATION: US/09/868,328

TIME: 14:07:32

Input Set : A:\EP.txt

Output Set: N:\CRF3\06202002\I868328.raw

128 IPSSDPNVYR YGGTLAPTIK LELVGHRSEV TEVKGYYIP PQIEWLAEDE SSNSAAALSL 420
E--> 129 LNK 423

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/868,328

DATE: 06/20/2002
TIME: 14:07:33

Input Set : A:\EP.txt

Output Set: N:\CRF3\06202002\I868328.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 9

VERIFICATION SUMMARY

DATE: 06/20/2002

PATENT APPLICATION: US/09/868,328

TIME: 14:07:33

Input Set : A:\EP.txt

Output Set: N:\CRF3\06202002\I868328.raw

L:30 M:283 W: Missing Blank Line separator, <400> field identifier
L:70 M:283 W: Missing Blank Line separator, <400> field identifier
L:79 M:252 E: No. of Seq. differs, <211> LENGTH:Input:521 Found:0 SEQ:2
L:85 M:283 W: Missing Blank Line separator, <400> field identifier
L:121 M:283 W: Missing Blank Line separator, <400> field identifier
L:129 M:252 E: No. of Seq. differs, <211> LENGTH:Input:423 Found:0 SEQ:4

<110> Smith, John; Smithgene Inc.

<120> Example of a Sequence Listing

<130> 01-00001

<140> PCT/EP98/00001
<141> 1998-12-31

<150> US 08/999,999
<151> 1997-10-15

<160> 4

<170> PatentIn version 2.0

<210> 1
<211> 389
<212> DNA
<213> Paramecium sp.

<220>
<221> CDS
<222> (279)...(389)

<300>
<301> Doc, Richard
<302> Isolation and Characterization of a Gene Encoding a
Protease from Paramecium sp.
<303> Journal of Genes
<304> 1
<305> 4
<306> 1-7
<307> 1988-06-31
<308> 123456
<309> 1988-06-31

<400> 1
agctgtagtc attcctgtgt cctcttctct ctgggcttct caccctgcta atcagatctc 60
aggagagagtg tcttgaccct cctctgcctt tgcagcttca caggcaggca ggcaggcagc 120
tgatgtggca attgctggca gtgccacagg cttttcagcc aggccttaggg tgggttcgcg 180
cgcggcgcg 1 cggccctct cgcgtctctc tcgcgcctct ctctcgctct cctctcgctc 240

Appendix 3, page 2

ggacctgatt aggtgagcag gaggagggggg cagtttagc atg gtt tca atg ttc agc 296
Met Val Ser Met Phe Ser

ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgt ttg ttc caa 344
Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val Cys Leu Phe Gln

tgt ccc aaa gtc ctc ccc tgt cac tca tca ctg cag ccg aat ctt 389
Cys Pro Lys Val Leu Pro Cys His Ser Ser Leu Gln Pro Asn Leu

<210> 2
<211> 37
<212> PRT
<213> Paramecium sp.

<400> 2
Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu
1 5 10 15

Phe Val Cys Leu Phe Gln Cys Pro Lys Val Leu Pro Cys His Ser Ser
20 25 30

Leu Gln Pro Asn Leu
35

<210> 3
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Designed peptide based on size and polarity to act as a linker between the alpha and beta chains of Protein XYZ.

<400> 3
Met Val Asn Leu Glu Pro Met His Thr Glu Ile
1 5 10

<210> 4
<400> 4
000

[Annex VIII follows]

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M, when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO: #:	Response shall be an integer representing the SEQ ID NO shown	M
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M

<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
<222>	Location	Specify location within sequence; where appropriate state number of first and last bases/amino acids	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified

base was used in
a sequence

Other relevant
information;
four lines maximum

M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.

Leave blank:
after <300>

Q

Preferably max
of ten named
authors of publi-
cation; specify
one name per line;
preferable format:
Surname, Other
Names and/or
Initials

0

Title

O

Journal

0

Volume

0

Issue

O

Pages

O

Date _____

Journal date on which
data published;
specify as yyyy-mm-
dd, MMM-yyyy or
Season-yyyy

O

Database
Accession
Number:

Accession number
assigned by data-
base including
database name

O

Database Entry
Date

Date of entry in
database; specify
as yyyy-mm-dd or
MMM-yyyy

0

Patent Document
Number

Document number;
for patent-type
citations only.
Specify as, for
example, US
07/999,999

0

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	0
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	0
<313>	Relevant Residues	FROM (position) TO (position)	0
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	0

5. Section 1.024 is revised to read as follows:

1.024 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by 1.021(c) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;